Measuring Representation: Perils of the Correlation Coefficient
Author(s): Christopher H. Achen
Source: *American Journal of Political Science*, Vol. 21, No. 4 (Nov., 1977), pp. 805-815
Published by: Midwest Political Science Association
Stable URL: http://www.jstor.org/stable/2110737
Accessed: 05/04/2013 18:06

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Midwest Political Science Association is collaborating with JSTOR to digitize, preserve and extend access to *American Journal of Political Science.*

http://www.jstor.org

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

CHRISTOPHER H. ACHEN
University of California, Berkeley

Measuring Representation: Perils of the Correlation Coefficient *

Most empirical research on representation has used correlational measures. The larger the correlation between representatives' and constituents' views, the stronger the bonds between them are thought to be. Unfortunately, correlations incorporate not only the strength of a relationship, but also the diversity of the sampled constituencies. When constituencies are very different from each other on an issue dimension, large correlations will result even when voters are not particularly sensitive to that dimension. Large correlations can occur when representatives are distant from their constituents; small correlations can happen when they are near. Correlations should be abandoned in the study of representation.

In liberal democratic theory, governments are meant to reflect the political values of their citizens. Popular preferences should control at least the fundamental direction of public policy. Not every transitory majority may get its way, and not every civic dispute may be settled democratically; but in principle, a liberal democracy makes the people sovereign. (See Mill, 1962, Ch. III; Dahl, 1956; Pitkin, 1967.)

Scholars have investigated the performance of contemporary Western democracies in meeting this standard. In a path-breaking essay of several years ago, Miller and Stokes assessed the "representativeness" of the 1958 American House of Representatives on three issue dimensions—social welfare, foreign policy, and civil rights. Using correlation coefficients, they found that in expressed policy preferences and in roll-call voting, Congressmen better reflected mean constituency opinion in civil rights matters than in social welfare or foreign policy (Miller and Stokes, 1966, p. 359).[1] Miller and Stokes also reported that winners of Congressional elections were more representative than losers on the social welfare dimension (Miller and Stokes, 1966; Miller, 1964).

*I am indebted to John Aldrich, David Mayhew, Hanna Pitkin, Edward Dorn, Peter Sperlich, Tom Moran, Gillian Dean, Lloyd Etheredge, Lynda W. Powell, Douglas Rae, and especially to G. Bingham Powell Jr., for helpful discussions and criticisms. The Inter-University Consortium for Political Research, Survey Research Center, University of Michigan, supplied the data. None of the above shares any responsibility for the errors of my analysis or conclusions.

[1] Miller and Stokes use Guttman scales as their measures of attitudes and roll calls.

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

In the decade following the original article, similar work has been carried out in other countries. Representatives and constituents have been interviewed and their responses compared in Austria, Italy, and Germany (Powell, 1974; Powell and Powell, 1975; Barnes and Farah, 1972), and limited cross-national comparisons are now possible. Together, these studies represent our best current wisdom on national legislatures as opinion-representing institutions.

Increasingly, however, certain conceptual difficulties are emerging in this line of research. "Opinion representativeness" has been measured by Pearson r's, tau-b's, tau-c's, gammas, and regression lines; and no consensus on their use has arisen. Attempted comparisons across countries or across issue dimensions soon run afoul of the crisscrossing, antagonistic assumptions and procedures employed in different studies. At best, one must speak of "strong" and "weak," "some" and "none"—precisely the vagueness statistical work is meant to cure.

The central difficulty is not statistical, but conceptual. Rarely is a measure of representativeness related to the ideas of liberal democratic theory—for example, citizen equality and popular sovereignty. Instead, measures have been plucked from the statistical shelf and employed without much theoretical interpretation. What aspect of liberal democracy does a gamma embody?

No statistical measure, of course, can contain the rich meaning that an idea like "popular sovereignty" has acquired in democratic theory. The humanistic tradition is simultaneously too rich and too vague. Statistical measures are not meant to replace those concepts, but instead to trade off some of their subtlety for an equally desirable analytical rigor and empirical precision. But needless to say, the latter benefits do not follow from *any* statistical procedures, but only from those with a demonstrable connection to theory. Without that connection, statistical work means little. To replace the unmeasurable by the unmeaningful is not progress.

The Correlation Coefficient

The most prominent statistics in the empirical representation literature are the correlation coefficient and its cousins. Miller and Stokes used Pearson r's, as did Barnes and Farah in their study of Italy and Germany. Powell employed a variant of the rank correlation coefficient (tau-C). All the researchers report trying a variety of correlational measures, but only Powell discusses statistical treatments of a noncorrelational sort. In sum, correlations are most of what we know statistically about representation.

Social scientists have long been aware of the weaknesses of the correlation

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

coefficient as a measure of strength in statistical relationships. (See Tufte, 1970, pp. 441–443.) Its worst property is that it reflects the particular measure and sample at hand as much as the underlying theoretical process. Thus correlations cannot be compared across samples in the way that unstandardized regression coefficients can: two correlations can differ because the variances in the samples differ, not because the underlying relationship has changed. For this reason, the "percent of the variance explained" tells us as much about the sample as about "explanatory power." Any single relationship can "explain" 5 percent or 95 percent of the variance, depending on the sample, while the theoretical process and predictive power remain the same.

The difficulty is illustrated in Figure 1. In both graphs, the same relationship is at work—a linear model in which $y = a + bx + u$, where u is an independently distributed disturbance term, uncorrelated with x, with mean zero and constant variance. In the right-hand sample, the point cloud is elongated and the R^2 is sizable; in the other graph, the values of the independent variable are not much dispersed and the R^2 is low. Yet the true model is precisely the same in each case: neither the "strength" of the independent variable nor the dispersion of the disturbances varies from graph to graph.

When we compare different samples, then, a small R^2 gives no guarantee that a relationship is weak, nor is a large R^2 a guarantee that it is strong. The same structure can yield any R^2 depending on the variances of the independent variables. Malinvaud's *Statistical Methods of Econometrics* (p. 90) characterizes the R^2 succinctly: "However, we must point out that in the present [regression] context, this coefficient is not particularly interesting. It gives no information about the goodness of the fit, but only about the shape of the collection of points."

Precisely the same objections apply to other correlational measures, like gamma or tau-b, and to standardized regression or path coefficients. (See Blalock, 1967; Duncan, 1975.) These quantities reflect both the "strength" of a variable and the diversity of its sampled values. If the sample or the variables are changed, the variances will usually change, too, and comparisons of explanatory power across samples or across variables will be impossible. The only comparable measures, and the quantities of genuine interest, are the unstandardized regression coefficients, along with the variance of the disturbances. These statistics are the theoretical parameters of the regression process, and unlike correlational measures, they reflect only the process at work, not the heterogeneity of the sample.

In representation studies, therefore, special danger accrues to the use of correlational measures. By their nature correlations destroy comparability

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

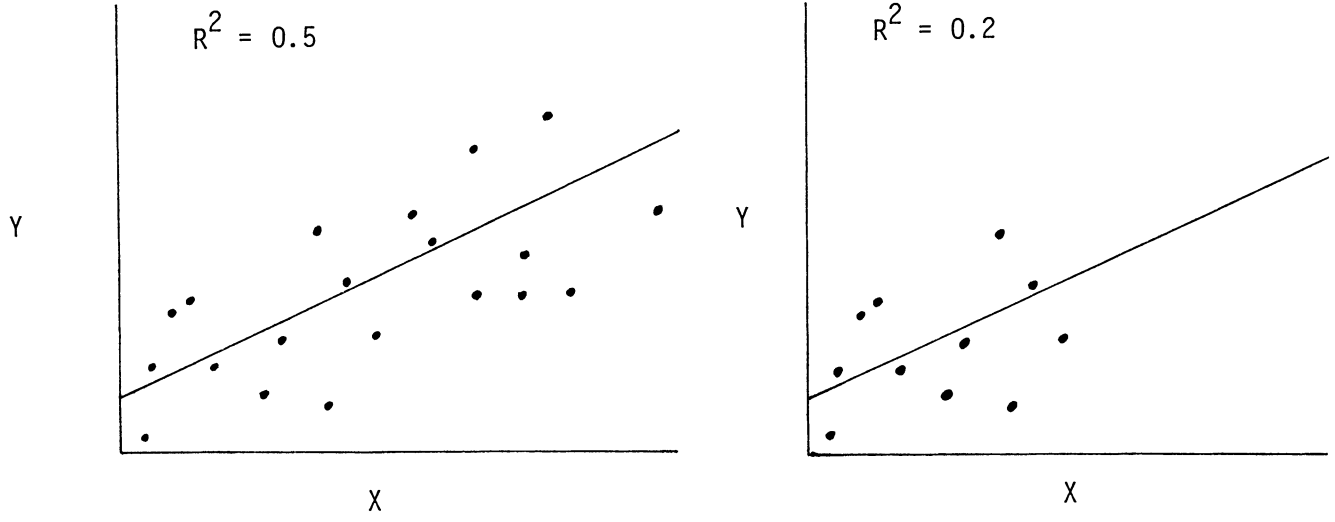

Figure 1

R^2 for the Same Relationship in Two Different Samples

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

across samples, meaning that representatives from two different countries cannot be compared nor can two different policy dimensions in a single country. Comparisons across policy dimensions or across nations will be infeasible. Large correlations simply do not demonstrate that representatives are closely associated with the views of their constituents.

The distortions caused by correlational measures are most easily illustrated for the Pearson r. Suppose that for some issue dimension, the views of winning candidates constitute a random sample from their constituencies. That is, the electoral system does not act to move candidates toward the center or toward any other position; instead, it acts like a lottery among the district's constituents. Candidates are produced who are equally likely to have the opinions of anyone in the district.

Intuition suffices to demonstrate that such a system will produce some correlation between candidates' views and the mean opinion of their districts. A random citizen of San Francisco will generally have more liberal views on civil rights issues than a random citizen from white, rural Alabama. Hence, randomly chosen representatives will typically differ from each other in the same way that their constituencies do, and a positive correlation between representative and constituency will be induced.

The size of this correlation can be estimated as follows. Suppose the true mean opinion in the jth constituency is u_j ($j = 1, \ldots, m$), and that the variance of opinion is σ_j^2. Denote the mean of the σ_j^2 by σ^2. Then a single randomly sampled observation from the jth constituency, x_j, may be written as $u_j + e_j$, where e_j is an error term independent of u_j with mean zero and variance σ_j^2. Finally, by u, e, and x we shall denote the random variables whose observed values are the u_j, e_j, and x_j.

By the usual formulas for the Pearson r^2, we have for the correlation between u and x:

$$r = \frac{\text{cov}^2(u,x)}{\text{var}(u)\text{var}(x)}$$

$$= \frac{\text{cov}^2(u, u+e)}{\text{var}(u)\text{var}(u+e)} \text{ (by definition of } x)$$

$$= \frac{\text{var}^2(u)}{\text{var}(u)[\text{var}(u) + \sigma^2]} \text{ (since } u \text{ and } e \text{ are uncorrelated)}$$

$$= \frac{\text{var}(u)}{\text{var}(u) + \sigma^2} \qquad (1)$$

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

In practice, of course, we do not observe the true means u_j, but must estimate them from a within-constituency sample as $\bar{x}_j$. The sampling error variance of $\bar{x}_j$ is of course σ_j^2/n_j, where n_j is the number of observations on which the sample mean is based in the constituency. We denote this latter quantity by ω_j^2, and its mean across the constituencies by ω^2. Now denote the variance of the $\bar{x}_j$ across the constituencies by $\text{var}(\bar{x})$. We have:

$$\text{var}(\bar{x}) = \text{var}(u) + \omega^2, \tag{2}$$

by the usual rule that the variance of a random variable observed with independent errors is equal to its systematic variance plus the error variance.

We are now in a position to estimate all the quantities in (1) from sample information. As in the Miller-Stokes case, suppose that we are given a set of m sample constituencies, each with n_j observations. The usual formulas for the mean and unbiased estimator of the variance yield $\bar{x}_j$ and $\hat{\sigma}_j^2$, our estimates of the mean and variance of opinion within the jth constituency. The same variance estimation formula applied to the $\bar{x}_j$ gives $\hat{\text{var}}(\bar{x}_j)$, the estimated variance of the $\bar{x}_j$ across constituencies; and ω_j^2 may be estimated as usual by $\hat{\omega}_j^2 = \hat{\sigma}_j^2/n_j$. Averaging the $\hat{\sigma}_j^2$ and $\hat{\omega}_j^2$ across constituencies gives estimates of σ^2 and ω^2, denoted by $\hat{\sigma}^2$ and $\hat{\omega}^2$, respectively. And finally, the estimate of $\text{var}(u_j)$ is: $\hat{\text{var}}(u) = \hat{\text{var}}(\bar{x}) - \hat{\omega}^2$ from (2). All these estimates are unbiased estimators.

Inserting the sample estimates above for $\text{var}(u)$ and σ^2 into (1), we obtain a simple estimate of the expected r^2 when sampling of representatives' opinions from constituencies is done randomly. The square root estimates r. For the Miller-Stokes data, we restrict ourselves to constituencies in which the winning candidate and five or more constituents were interviewed. For those constituencies, the results in Table 1 are obtained.[2]

[2] The observed correlations in my sample, corrected for measurement error, were: social welfare, 0.16; foreign policy, 0.49; civil rights, 0.46. The difference between these correlations and the observed correlations in Table 1 can probably to attributed to (1) Miller and Stokes' use of weighting, and (2) differences in the definition of the sample. However, Miller and Stokes apparently reserved much of the technical details of their procedures to a never-completed book–hence, it is difficult to be certain. In any case, Pearson r's are notoriously unstable, and the differences are not surprising.

If the sample is restricted to Northerners, then the expected correlations are (with the observed in parentheses): social welfare, 0.28 (0.19); foreign policy, 0.00 (0.18); and civil rights, 0.22 (0.14). None of the observed correlations approaches a statistically significant difference from its expectation.

Robert Erikson has reminded me that the corrections for measurement error used here depend on the assumption of simple random sampling, and therefore the observed

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

TABLE 1

Correlation between Constituency Opinion and Views of Representatives

	Social Welfare	Foreign Policy	Civil Rights
Expected Pearson r[a]	0.27	0.16	0.35
Observed Pearson r[b]	0.26	0.32	0.50

[a]Estimates of the expected correlation (r) between true constituency mean opinion and views of a hypothetical representative randomly selected from the constituency for the Miller-Stokes Congressional districts where the representative and at least five respondents were interviewed.

[b]Observed correlation (r) between actual representatives' views and constituency mean opinion, corrected for measurement error, as reported by Miller and Stokes (1966).

These correlations show that civil rights views correlate more strongly with constituency opinion than do Congressmen's views on the other two issue dimensions. And yet we have assumed that the process at work was the same on each dimension: the electoral system produces a randomly selected citizen from the district. In short, *correlations with the Miller-Stokes pattern can be obtained from an electoral system whose constituents ignore representatives' opinions.*

Similar arguments can be made for correlations between Congressmen's roll-call voting and constituency opinion or between perceptions of district opinion and true constituency views. Positive correlations between these variables, larger for civil rights than for the other two dimensions, can be generated by random mechanisms without assuming that Congressmen or constituents know anything at all about each other. Gammas, rank correlation coefficients, and other correlational measures will behave in essentially the same way.

Random selection of representatives from a district is not entirely undemocratic, of course, and a measure of representativeness should reflect that. A random system is surely preferable to representatives of a single viewpoint who are imposed on the constituencies from outside or even to representatives chosen at random from the nation as a whole. These latter methods would produce correlations of zero between representative and

r's, corrected for measurement error, should be even larger. It is also true that taking account of the nature of the sample would reduce the expected r's somewhat. However, these adjustments to both observed and expected correlations are probably quite modest.

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

constituents; random selection within districts will necessarily do better. The correlations in Table 1 *should* be positive.

The objection to correlations in this instance is not that they produce positive values, but rather that the correlations fail to capture the underlying notion that representatives should be close to the center of their districts. The correlations in Table 1 misrepresent the proximity of the representatives to their citizens. When selection is random within constituencies, the proximity of a representative depends on the within-constituency variance. The more homogeneous the district, the better a randomly selected Congressman will be as a representative of average opinion. Ironically, in the Miller-Stokes data, it is social welfare that has the smallest within-constituency variance. In fact, in the analysis for which correlations are reported in the first row of Table 1, citizens turn out to be *closest to their randomly chosen representative in social welfare opinion, not civil rights.* (Let each opinion scale be one unit long. Then the expected squared difference between a randomly selected representative and a constituent is: social welfare, 0.264; foreign policy, 0.401; civil rights, 0.289.) Thus even the modest form of representation resulting from random selection escapes measurement by the correlation coefficient.[3]

These misleading correlations have their source in the very different dispersions of mean constituency opinion on the three dimensions. The South differs greatly from the North on civil rights, much less so on social welfare and foreign policy. As usual, the more diverse the independent variable, the larger the correlation, so that civil rights correlations cannot help but be larger than those for social welfare and foreign policy even when the regression parameters are the same. The iron law of correlation coefficients is that the sizes of correlations from different samples do not necessarily reflect

[3] For all three dimensions in Table 1, the unstandardized regression coefficients, for the random selection case corrected for measurement error, are a slope of one and an intercept of zero. By this measure of "representativeness," all three dimensions are equally representative under random sampling of Congressmen. Thus the correlations are misleading by this definition as well.

Note, too, that the standardized betas in this example are the same as the correlations, and therefore larger for civil rights than for the other two dimensions. Yet the true slopes are identical on all three dimensions, and citizens are actually closer to their representative on foreign policy and social welfare, illustrating how standardization misrepresents the underlying mechanisms. Betas reflect both the "strength" of a variable and the diversity of its sampled values, and are therefore a particularly unfortunate choice of scale for variables with no natural units if a researcher wants to make comparisons across samples.

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

the relative strengths of the underlying relationships. The larger variance of mean constituency civil rights opinion makes the r's virtually useless as guides to inference about representation.[4]

Conclusion

The perennial difficulty of giving statistical interpretation to social science theory occurs in acute form in the study of representation. The conceptual subtleties of the liberal theory of republican democracy have been largely ignored by empirical researchers, and ready-to-hand statistical measures within little or no substantive meaning have served to summarize large bodies of data. Most of these measures have been correlational, with all the attendant weaknesses of such statistics. When used to study representation, correlations mix together the strength of the representative-to-constituent connection with the diversity of the constituencies. In the Miller-Stokes data, the district heterogeneity is so much larger for civil rights than for the other two dimensions that a larger correlation for that dimension is almost inevitable. In fact, the pattern of the Miller-Stokes correlations can be produced from a system of random selection of representatives within districts. The larger correlation for civil rights, then, does not imply that constituency opinions about race influence Congressmen more than any other issue dimension. Correlations are simply incomparable across different variables.

In principle, correlational measures ought to be replaced by unstandardized regression coefficients plus the variance of the residuals. For the Miller-Stokes data, regression would produce estimates of the influence of constituents on representatives, and those estimates would be comparable across dimensions. Given a constituency mean opinion, the expected position of the representative on the same dimension could be predicted, and the average error of the prediction could be estimated as well.

Unfortunately, in the study of representation, nothing is quite so easy. First of all, the regression estimates cannot be produced simply by regressing the representative's opinion on the sample mean of his constituents' views. Any data set for representation necessarily includes only a selection of citizen opinion, so that true constituency mean opinion is not known with certainty.

[4] Of course, one could choose to *define* "representativeness" as the Pearson r. The point here is that such a definition is likely to be difficult to work with and misleading in its implications. It should be added that Miller and Stokes noted the possibility of correlations arising from random sampling (p. 362), but they seem to have overlooked its dangers for their analysis.

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

In the Miller-Stokes case, some constituencies are represented by as few as half a dozen interviews. The true mean is unknown. This measurement error in the independent variable can bias regression coefficients severely, and special procedures must be adopted to obtain useful estimates.

Moreover, even if attractive regression estimates could be obtained, the problem of interpretability remains. Substituting regression for correlation surely represents a technical improvement, but what aspects of democratic theory are captured by regression parameters? Without a connection to traditional thinking on representation, no statistical summaries of republican system are likely to be satisfactory. At present, no such connection exists, and without it, regression holds no special charm.

The statistical and theoretical aspects of the empirical study of representation pose severe conceptual difficulties for the student of democracy. No simple solutions will be forthcoming; certainly none will be attempted here. But even the modest results presented above—the frailties of the correlation coefficient—should serve to remind us that no data are ever naked facts: they arrive packaged in the concepts and categories we have used to collect them. This is nowhere more true than in quantitative work, where large bodies of information are often reduced to a few measures. All the more reason, then, that this initial stage of theory-building be given the closest attention, so that researchers can at least be certain what it is they must explain. Unlike those fields of knowledge in which experimentation is commonplace, the social sciences must often expend considerable analytical effort simply to establish basic empirical relationships. If this stage is slighted, for example by choosing ready-to-hand statistical measures with little or no substantive meaning, the result will inevitably be faulty versions of the facts and misguided attempts to explain them.

Manuscript submitted 9 June 1976
Final manuscript received 29 March 1977

REFERENCES

Barnes, Samuel H. and Barbara Farah. 1972. National representatives and constituency attitudes in Germany and Italy. Paper prepared for delivery at the 1972 Annual Meeting of the American Political Science Association. Washington, D.C., September 5–9.

Blalock, H. M. Jr. 1967. Causal inferences, closed populations, and measures of association. *American Political Science Review,* 61 (March 1967): 130–136.

Dahl, Robert A. 1956. *Preface to democratic theory.* Chicago: University of Chicago Press.

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions

Duncan, Otis Dudley. 1975. *Introduction to structural equation models.* New York: Academic Press.

Malinvaud, E. 1970. *Statistical methods in econometrics,* 2nd revised edition. New York: American Elsevier.

Mill, John Stuart. 1962. *Considerations on representative government.* Chicago: Henry Regnery.

Miller, Warren E. 1964. Majority rule and the representative system of government. In Erik Allardt and Yrjo Littunen, *Cleavages, ideologies and party systems.* Helsinki: Academic Bookstore, pp. 343–76.

____ and Donald E. Stokes. 1966. Constituency influence in Congress. In Angus Campbell, Philip E. Converse, Warren E. Miller, and Donald E. Stokes, *Elections and the political order.* New York: Wiley, pp. 351–373.

Pitkin, Hannah. 1967. *The concept of representation.* Berkeley: University of California Press.

Powell, G. Bingham Jr. 1974. Citizen-elite linkages in Austrian communities: Degree, structure and process of opinion representation. Paper prepared for delivery at the 1974 Annual Meeting of the American Political Science Association, Chicago, August 29–September 2.

____ and Lynda W. Powell. 1975. The analysis of citizen-elite linkages. Unpublished manuscript, July 1975.

Tufte, Edward R. 1970. Improving data analysis in Political Science. In Tufte, ed., *Quantitative analysis of social problems.* Reading, Mass.: Addison-Wesley.

This content downloaded from 130.192.119.156 on Fri, 5 Apr 2013 18:06:07 PM
All use subject to JSTOR Terms and Conditions